Cepstral/Spectral Index of Dysphonia Estimates and Perceptual Ratings
of Voice in Adolescents with Supraglottic Phonation Following
Pediatric Airway Reconstruction

Alessandro de Alarcon MD, MPH; Lisa N. Kelchner, Ph.D.;
Barbara Weinrich, PhD.; Robert E. Hillman, PhD.; Dimitar Deliyski, PhD.

Results:

Table 1. Demographics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age</th>
<th>Number of Previous Airway Surgeries</th>
<th>Source of Vibration for Voice*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>18</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>15</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>18</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>15</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>16</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>18</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

*Legend for Vibration Source: 1= combined or multiple supraglottic structures; 2= ventricular folds

Figure 1. Easy Onset Sentence: Graphical Display of CSID Estimate and CAPE-V Overall Severity Ratings for Raters 1 & 2

Figure 2. All Voiced Sentence: Graphical Display of CSID Estimate and CAPE-V Overall Severity Ratings for Raters 1 & 2

Figure 3. Easy Onset Sentence: Graphical Display of CSID Estimate and CAPE-V Roughness Rating For Raters 1 & 2

Figure 4. Easy Onset Sentence: Graphical Display of CSID Estimate and CAPE-V Breathiness Rating For Raters 1 & 2

Summary

- Participants 1-4 had a greater degree of roughness perceived by the raters; the ADSV software generated estimates for those participants that were lower. In some cases the voice signal, while "rough sounding", was more periodic despite the altered vibration source and as such the derived estimate of dysphonia was low. The ADSV program does not factor age/gender considerations into the estimate which is often defining parameter of subjective-perceptual ratings of voice made by clinicians.
- There was a notable improvement in the similarity of the ratings and CSID estimates for the parameter of breathiness--suggesting stronger "agreement" between the listener raters and the ADSV program.
- The ADSV shows promise as an alternative tool for acoustic analysis of voice quality for patients with severe dysphonias and supraglottic phonation. When comparing to clinical raters, care has to be taken if roughness comprises the majority of overall severity.

Citations


Materials and Methods

Participants

- Ten children (6 male; 4 female), ages 6-18 years (median: 15 years) post-airway reconstruction with previously identified severe dysphonia and/or alternate source of vibration for voice were selected for inclusion in this study from a large pediatric voice registry database. Demographic data including source of vibration for voicing for each participant are shown in Table 1.

Analysis of Dysphonia in Speech and Voice

- Using the Real-Time Pitch program of the Computerized Speech Lab (KayPentax), all voice recordings were collected in a sound proof booth, controlling for microphone type, microphone-mouth distance and sampling rate per our established CCHMC-IRB approved voice registry and voice clinic protocol.
- Two CAPE-V sentences, How hard did he hit him (easy onset) and We were away a year ago (all voiced) were selected from each participant’s file and edited and then analyzed per ADSV guidelines for analysis. Because of the perceived severity of many of the voice samples, the Cepstral Peak Prominence threshold and display range minimum were set at -5dB.
- The Cepstral Peak Prominence in dB (CPP (dB)), Cepstral Peak Prominence standard deviation in dB (CPP SD (dB)), Low/High frequency spectral energy ratio (L/H ratio), and Low/High frequency spectral energy ratio standard deviation (L/H ratio-SD), and the CSID estimate were calculated and entered into an excel file.

Expert Perceptual Ratings

- Using the CAPE-V, new individual expert perceptual ratings of Overall Severity (OS), Roughness (R) and Breathiness (B) were completed for each of the two sentences for a total of 60 ratings per rater.
- Each rater re-rated 3 samples at a later time.

Data Analysis:

In order to inspect expert rating patterns and compare them to CSID estimates, line graphs were generated from the rating and estimate data in the excel database. Data were sorted by rater and sentence type. We examined the findings for CSID estimates and CAPE-V OS (for each sentence type), roughness, and breathiness. Summary data were tabulated.